DefPol: New Procedure to Build Molecular Surfaces and Its Use in Continuum Solvation Methods

CHRISTIAN S. POMELLI,[1] JACOPO TOMASI[2]

[1]*Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy*
[2]*Dipartimento di Chimica e Chimica Industriale, Via Risorgimento 35, 56124 Pisa, Italy*

Received 9 February 1998; accepted 13 July 1998

ABSTRACT: We present a method to define van der Waals, solvent-accessible, and solvent-excluding molecular surfaces with their partition in nonoverlapping surface portions (tesserae). The procedure is more efficient than those available in the literature to describe solvent effects on molecular systems of large size, and it can also be applied to solutes of small size without reducing the accuracy of the output and without increasing computational times. All the tesserae are expressed in terms of spherical triangles, having all the characterizing elements (vertices, centers, etc.) analytically defined. The method was tested by comparing the results for the surface area and the solvation free energy (decomposed in electrostatic, dispersion, and steric contributions) obtained using the GEPOL procedure within the framework of the polarizable continuum model solvation method. These comparisons regard 87 molecules at the molecular mechanics level and 28 molecules at the *ab initio* Hartree–Fock level: the results are quite satisfactory. © 1998 John Wiley & Sons, Inc. J Comput Chem 19: 1758–1776, 1998

Keywords: DefPol; molecular surfaces; continuum solvation methods

Introduction

Continuum methods addressing the study of solvent effects on molecular systems rely heavily on the concept of the molecular cavity in the continuum solvent distribution in which the solute is accommodated. To have a realistic description of solvent effects within this approach, the cavity must be accurately modeled on the shape of the molecule. This requirement introduces a strict correlation between the cavity and the molecular volume: the cavity can be viewed as the mold of the solute in this continuum solvent distribution.

Correspondence to: J. Tomasi; e-mail: tomasi@cl.c.ci.unipi.it
Contract/grant sponsors: CNR, MURST

CCC 0192-8651 / 98 / 151758-19

Among the several approaches to continuum solvation models,[1] the most versatile, efficient, and precise ones are based on the use of surface effectors. The most important of such effectors are the apparent surface charges (ASCs) describing the electrostatic reaction field, but there are similar effectors to set the solvation energy for the other contributions to the solvation energy.[1] The use of these effectors requires a partition of the cavity surface into relatively small portions called tesserae; these cover all the cavity surface without overlaps, with a well defined area, and are complemented by other geometrical information, such as the location of some representative points. These geometrical data must be more detailed when the computational algorithm also addresses the evaluation of the equilibrium geometries of the solute and the transition state in reactions; both quantities require the analytical evaluation of energy derivatives with respect to the nuclear coordinates.

The use of molecular volumes and surfaces is not limited to the solvation continuum methods. The volume and surface of a molecule are concepts without a clear cut quantum definition but are of widespread use in chemistry[2,3] and often require accurate and detailed calculations, accompanied by a partition of the surface into small portions (an example is the use of molecular surfaces in computational graphics).

A large number of articles concerning algorithms to compute molecular surface and volumes can be found in the literature.[4] Some of these methods lead to a partitioned surface; some give surface area and volume values only.

A vivid review by Connolly on molecular surfaces can be found on the worldwide web.[5] This review presents a general overview and provides a large bibliography on the various definitions, algorithms, and uses of molecular surfaces.

A discussion about the numerous algorithms thus far proposed to compute molecular surfaces is beyond the scope of this article. We refer to Connolly's review for more exhaustive bibliographic references. We focus on those more largely used: Connolly's dot method[6,7] and the GEPOL procedures.[8,9]

Connolly's method is largely used in computer graphics and it has also been used in solvent effects studies: the number of dots in its standard version is relatively large to allow a good graphical representation; this fact may represent a disadvantage in solvent effects calculations that may be computer expensive when the method is applied to large molecules. In addition, there are some difficulties in using this description to compute analytical derivatives of the solvation energy. (However, see Zauhlar and Morgan.[10,11])

GEPOL uses less tesserae for unit area: this is detrimental to the graphical rendering, but we showed[12] that the surface partitioning scheme used in the standard GEPOL is the most efficient one to describe solvent effects, using as parameters for the comparison the number of tesserae one uses to reach a given accuracy. Ahuir and Silla also showed on several occasions that GEPOL is more efficient than Connolly's algorithm in describing the volume and surface of complex molecules.[13]

GEPOL was originally conceived as part of the polarizable continuum model (PCM),[14] a solvation method based on the use of surface effectors such as the ASC we mentioned; and it has been used in all the numerous PCM-derived methods. The quality of this description recently led to the use of these algorithms in other continuum solvation methods that are not based on ASCs, which in preceding versions used other definitions of the cavity surface.[15,16] It may be profitable to use for calculations of geometrical derivatives of the energy, and there are now several versions of geometry optimization programs exploiting GEPOL.[17]

This method has some shortcomings, however, especially when the molecule is large: it depends on the basic strategy used to define the volume and cannot be easily eliminated if this basic strategy is maintained.

Therefore, there is a need for an alternative strategy that is able to eliminate these shortcomings without losing important features of GEPOL, such as the analytical definition of the necessary representative points of each tesserae, the efficiency in the evaluation of solvent effects, and the easy insertion in automated derivative calculations.

This study focuses on a new procedure for the calculation of the cavity (or molecular) surface with the above-mentioned characteristics.

Problem

The molecular volume in GEPOL is described in terms of the union of van der Waals spheres centered on the atoms. (For the calculation of the electrostatic contributions to the solvation energy it has been found convenient[1] to enlarge the spheres by multiplying the van der Waals radii R_k by a factor f_k near 1.2.) Each sphere is partitioned

into tesserae defined in terms of an inscribed pentakisdodecahedron (60 triangular tesserae with the same shape). The molecular surface, $\mathscr{S}_w$, is defined in terms of spherical surfaces portions, each portion being partitioned in a number t_k of tesserae that are mostly of triangular shape and with area related to $f_k R_k$ and some of irregular shape at the junction between or among spheres. To achieve analytical derivatives, GEPOL was refined with the inclusion of analytical calculations of the vertices of the tesserae (triangular and irregular).[18]

The use of the $\mathscr{S}_w$ surface is not sufficient in the field of solvation models. Two other cavities with surface $\mathscr{S}_a$ (ref. 19) and $\mathscr{S}_e$ (ref. 20) were introduced: GEPOL contains algorithms for their construction.

Lee and Richards[19] called the $\mathscr{S}_a$ the ''accessible surface;'' it corresponds to a cavity where the centers of solvent molecules described as hard spheres of radius R_s are not allowed. It may be defined as the surface spanned by the center of a solvent sphere rolling on $\mathscr{S}_w$. The corresponding cavity may be described in terms of atomic spheres with radii $R_{vdW} + R_s$. The partition of $\mathscr{S}_a$ into tesserae is given in GEPOL as for $\mathscr{S}_w$.

The computational definition of $\mathscr{S}_e$ (solvent-excluding surface according to the terminology introduced by Richards[20]) is more delicate. By definition, it corresponds to a cavity where solvent molecules described as rigid spheres are not allowed. In smooth regions of the molecule $\mathscr{S}_e$ coincides with $\mathscr{S}_w$, but there are deviations where the molecule presents crevices or indents that are not large enough to accommodate a solvent molecule.

In GEPOL the space included within $\mathscr{S}_e$ but not included within $\mathscr{S}_w$ is described by filling it with spheres with different radii. The process is fully automated with a progressive filling of the space. In 1986 we selected this definition to have a spherical convex surface everywhere, for which the boundary element method (BEM) calculation of the electrostatic solvation energy was considered to be less subject to errors.[1] These additional spheres are subjected in standard GEPOL to the partition into tesserae as are the original ones.

In conclusion, the GEPOL procedure leads to additional spheres (and tesserae) for $\mathscr{S}_e$, the number of which in some cases, depending on the topography of the solute, can be quite large.[12]

The new procedure, DefPol, uses a different strategy. First a set of shape functions is defined. They describe the portion of space defined by the $\mathscr{S}_w$, $\mathscr{S}_a$, and $\mathscr{S}_e$ surfaces we described. Then each shape function is surrounded by a single ellipsoid that is large enough to contain the whole shape function. The ellipsoid has the main axes proportional to the inertia axes of the shape function and derives from a sphere suitably stretched in the three directions. The original sphere is partitioned into tesserae according to a ''regular'' partition into triangles that are given as input (60 or 240 faces or more), and consequently the ellipsoid is partitioned into stretched triangles. Each triangle is then ''moved'' to be better accommodated on the shape function and then further partitioned to achieve a better description of the surface.

In conclusion, all the three surfaces are described with the same algorithm in terms of triangular spheric surfaces with vertices, center, and radius of curvature analytically definite.

The potential advantages we found in defining this new strategy are summarized as follows:

1. to avoid an explosion in the number of GEPOL parameters: the number of additional spheres for $\mathscr{S}_e$ and the number of tesserae. For large molecules with irregular shape GEPOL gives an exceeding large number of tesserae.
2. to always have triangular tesserae: this feature simplifies the calculation of analytical derivatives with respect to nuclear coordinates. A limited number of tesserae can be concave, but the recent BEM formulations of the electrostatic problem[17] can handle a mixing of concave and convex tesserae without problems.
3. to monitor the number of tesserae involved: the number of tesserae given as input can be slightly enlarged during the refinement of the procedure, but there are no potential sources of ''explosion'' in the number of parameters.
4. to avoid the occurrence of spurious ''solvation pockets'': in large molecules void portions of the inner space with dimensions sufficient to accommodate one or more solvent molecules may be present. They are empty, but GEPOL is not able to detect this fact and introduces spurious additional solvation contributions.

In the following sections we document this procedure that, to the best of our knowledge, is completely original.

The documentation is organized as follows: we introduce the shape function and the algorithms for its calculation, describe the DefPol definition of the molecular surface and of its partition, present numerical results and the comparison with those obtained with GEPOL, and present a short final discussion.

Shape Function

DEFINITIONS

We define the portion of space occupied by a molecule through a *shape function*:

$$\mathcal{M}(\vec{p}) = \begin{cases} 1 & \vec{p} \text{ inside the molecule,} \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

To give an explicit form to the shape function we have to define a model of the molecular surface. In the van der Waals model a sphere of appropriate radius is centered on each atom. We define the atomic shape function as

$$\mathcal{M}^i_{\mathcal{S}_w}(\vec{p}) = \begin{cases} 1 & |\vec{p} - \vec{p}_i| \le R_i \\ 0 & \text{otherwise,} \end{cases} \quad (2)$$

where $\vec{p}_i$ and R_i are the coordinate and the radius of the ith atom. We write the van der Waals shape function for N atomic molecules as a combination of atomic shape functions:

$$\mathcal{M}_{\mathcal{S}_w}(\vec{p}) = \bigvee_{i=1}^{N} \mathcal{M}^i_{\mathcal{S}_w}(\vec{p}). \quad (3)$$

The symbol $\vee$ indicates that the various atomic terms are combined by an OR logical operator: $\mathcal{M}_{\mathcal{S}_w}(\vec{p})$ is not equal to zero if almost one of the $\mathcal{M}^i_{\mathcal{S}_w}(\vec{p})$ are not equal to zero. In the computational practice we evaluate all the atomic shape functions until one of these returns a value of one. If all the atomic shape functions are equal to zero, we are outside the molecule.

We can also describe by the appropriate shape function the solvent-accessible and the solvent-excluding surfaces. For the solvent-accessible surface we have:

$$\mathcal{M}^i_{\mathcal{S}_a}(\vec{p}) = \begin{cases} 1 & |\vec{p} - \vec{p}_i| \le R_i + R_s \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

where R_s is the solvent radius and

$$\mathcal{M}_{\mathcal{S}_a}(\vec{p}) = \bigvee_{i=1}^{N} \mathcal{M}^i_{\mathcal{S}_a}(\vec{p}). \quad (5)$$

The shape function for $\mathcal{S}_e$ has a more complex form. We have an atomic part

$$\mathcal{M}^i_{\mathcal{S}_e}(\vec{p}) = \begin{cases} 1 & |\vec{p} - \vec{p}_i| \le f_i R_i, \\ 0 & \text{otherwise,} \end{cases} \quad (6)$$

where f_i is a scale factor usually in the range of 0.90–1.25 and a multiatomic part that takes account of the steric interactions between the solvent and two or more of the solute's atoms. To have these steric interactions the spherical solvent probe must be tangent to all the atoms involved. The tangency condition is

$$|\vec{p}_i - \vec{p}_s| = R_i + R_s, \quad (7)$$

where $\vec{p}_s$ is the position of the solvent sphere. Each tangency condition represents a local constraint on the position of the probe sphere:

1. The solvent sphere is tangent to one atomic sphere. Locally there are no portions of the outer space not accessible to the solvent.
2. The solvent sphere is tangent to two atomic spheres. To have this situation the condition

$$|\vec{p}_i - \vec{p}_j| \le R_i + R_j + 2R_s \quad (8)$$

must be satisfied. The system has $C_{\infty v}$ symmetry; the probe sphere in its motion describes a thorus. We indicate this thorus with τ_{ij}. The space between the thorus and the two atomic spheres is not accessible to the solvent. The contribution of these regions of space to the shape function are defined by the symbol $\mathcal{M}^{ij}_{\mathcal{S}_e}(\vec{p})$.

3. The solvent sphere is tangent to three atomic spheres. To have this situation the system

$$\begin{cases} |\vec{p} - \vec{p}_i| = R_i + R_s, \\ |\vec{p} - \vec{p}_j| = R_k + R_s, \\ |\vec{p} - \vec{p}_k| = R_j + R_s \end{cases} \quad (9)$$

must have real solutions and the three atoms cannot be aligned. From a geometrical point of view, the probe sphere can occupy two positions that are specular with respect to the plane determined by $\vec{p}_i$, $\vec{p}_j$, and $\vec{p}_k$; thus, the system of eq. (9) subtends the C_s symmetry group. The centers of these spheres can be

found by solving system (9): we indicate them with σ^1_{ijk} and σ^2_{ijk}. The space between the three atomic spheres, σ^1_{ijk}, σ^2_{ijk}, and the regions where at least one of the terms $\mathcal{M}^{ij}_{\mathcal{S}_e}(\vec{p})$, $\mathcal{M}^{ik}_{\mathcal{S}_e}(\vec{p})$, $\mathcal{M}^{jk}_{\mathcal{S}_e}(\vec{p})$ is not equal to zero is not accessible to the solvent. The symbol $\mathcal{M}^{ijk}_{\mathcal{S}_e}(\vec{p})$ defines the contribution of these regions of space to the shape function.

4. The solvent sphere is tangent to four or more atomic spheres. If we apply four or more tangency constraints [eq. (7)], generally no allowed positions for the probe sphere are found. If only two or three of these constraints are linearly independent, we have to take these situations into account with the aid of two and three atom terms.

These considerations led to the following form of the $\mathcal{S}_e$ shape function:

$$\mathcal{M}_{\mathcal{S}_e}(\vec{p}) = \left(\bigvee_{i=1}^{N} \mathcal{M}^{i}_{\mathcal{S}_e}(\vec{p}) \right) \vee \left(\bigvee_{i<j}^{N} \mathcal{M}^{ij}_{\mathcal{S}_e}(\vec{p}) \right) \vee \left(\bigvee_{i<j<k}^{N} \mathcal{M}^{ijk}_{\mathcal{S}_e}(\vec{p}) \right). \quad (10)$$

Some of the $\mathcal{M}^{ij}_{\mathcal{S}_e}(\vec{p})$ and $\mathcal{M}^{ijk}_{\mathcal{S}_e}(\vec{p})$ terms are equal to zero in all the space because the atoms do not fulfill the conditions given in these components.

TWO ATOM TERMS

For the two atom component $\mathcal{M}^{ij}_{\mathcal{S}_e}(\vec{p})$ let us consider Figure 1: the symmetry of the system allows us to perform all the calculations in two dimensions. We indicate the two atoms i and j with the letters A and C and the centers of the two circular sections of τ_{ij} with B and D. Some simple geometrical considerations led to the following equations:

$$\begin{aligned} \overline{AG} &= \overline{AI} = R_i, \\ \overline{CH} &= \overline{CL} = R_j, \\ \overline{BH} &= \overline{BG} = \overline{DL} = \overline{DI} = R_S, \end{aligned} \quad (11)$$

where R_i, R_j, and R_S are the radii of atoms i, j, and the solvent sphere, respectively. The thorus τ_{ij} is determined by its center, arm, and radius and by the unit vector orthogonal to its symmetry plane. (For the definitions of these quantities, see Fig. 2.) The radius of τ_{ij} is simply the radius of the sphere that generates the figure; thus,

$$r_{\tau_{ij}} = R_S. \quad (12)$$

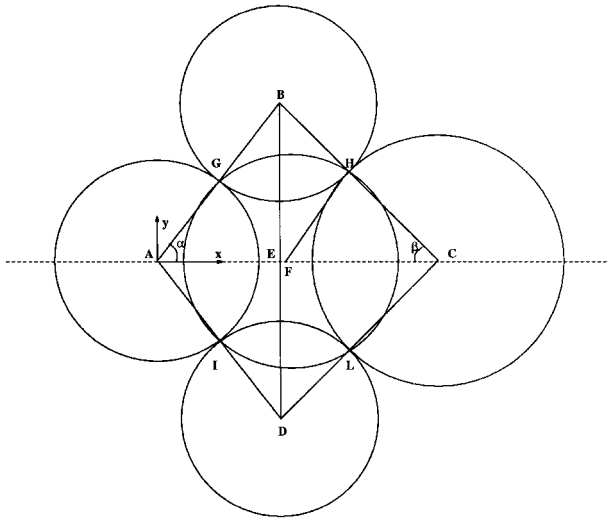

FIGURE 1. Geometrical construction used to determine the parameters of the two atom terms.

The unit orthogonal vector $\hat{n}_{\tau_{ij}}$ is

$$\hat{n}_{\tau_{ij}} = \frac{\vec{p}_j - \vec{p}_i}{R_{ij}}, \quad (13)$$

where $\vec{p}_i$ and $\vec{p}_j$ are the positions of the two atoms and $R_{ij} = \overline{AC}$ is the distance between them.

The center and the arm of the thorus are

$$\vec{c}_{\tau_{ij}} = \vec{p}_i + \overline{AE}\left(\vec{p}_j - \vec{p}_i\right) \quad (14)$$

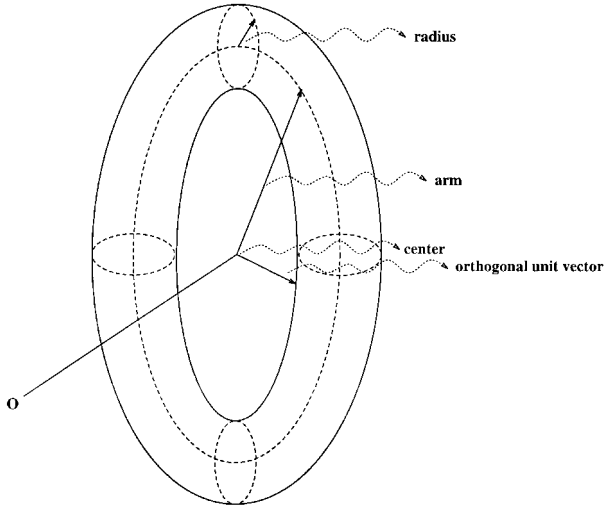

FIGURE 2. Pictorial definition of thorus-related quantities.

and

$$a_{\tau_{ij}} = \overline{BE}. \tag{15}$$

To obtain the lengths of $\overline{AE}$ and $\overline{BE}$ we have to apply the Pythagorean relation to the triangles AEB, BEC:

$$\overline{AE^2} + \overline{BE^2} = \overline{AB^2} = (R_i + R_S)^2, \tag{16}$$

$$\overline{EC^2} + \overline{BE^2} = \overline{BC^2} = (R_j + R_S)^2, \tag{17}$$

and observe that

$$\overline{AE} + \overline{EC} = R_{ij}. \tag{18}$$

After some simple algebraic manipulations we find that

$$\overline{AE} = \frac{R_{ij}^2 + (R_i + R_S)^2 - (R_j + R_S)^2}{2R_{ij}} \tag{19}$$

and

$$\overline{BE} = \sqrt{+(R_i + R_S)^2 - \overline{AE^2}}\,. \tag{20}$$

To evaluate $\mathscr{M}_{\mathscr{S}_e}^{ij}(\vec{p})$ we also need the center and the radius of the sphere circumscribed to the portion of space in which $\mathscr{M}_{\mathscr{S}_e}^{ij}(\vec{p}) \neq 0$ but $\mathscr{M}_{\mathscr{S}_e}^{i} = \mathscr{M}_{\mathscr{S}_e}^{j} = 0$. This portion of space is generated by the rotation of the curvilinear quadrilateral $GHLI$ of Figure 1 around the principal axis of symmetry of the system. We indicate this sphere with ω_{ij}. To obtain the center and the radius of this sphere, again consider Figure 1: the point F is the center of the sphere and of the section of the sphere in this plane. If we know $\overline{AF}$ and $\overline{FH}$ we also know the center and the radius of this sphere:

$$R_{\omega_{ij}} = \vec{r}_i + \overline{AF}\hat{n}_{\tau_{ij}} \tag{21}$$

and

$$\vec{p}_{\omega_{ij}} = \overline{FH}. \tag{22}$$

Consider now the coordinate system defined in Figure 1. To find the equation of the circular section of the sphere ω_{ij} we have to solve the system

$$\begin{cases} (X_G - X_F)^2 + Y_G^2 = R_{\omega_{ij}}^2, \\ (X_H - X_F)^2 + Y_H^2 = R_{\omega_{ij}}^2, \end{cases} \tag{23}$$

thus,

$$X_F = \frac{X_G^2 + Y_G^2 - X_H^2 - Y_H^2}{2(X_G - X_H)}, \tag{24}$$

$$R_{\omega_{ij}} = \sqrt{X_G^2 + Y_G^2 - 2X_G X_F + X_F^2}\,. \tag{25}$$

To find the coordinates of the points G, H, consider the triangles ABE and BCE:

$$X_G = \overline{AB}\cos\alpha = \pm\frac{R_i}{R_i + R_S}\overline{AE}, \tag{26}$$

$$Y_G = \overline{AB}\sin\alpha = \frac{R_i}{R_i + R_S}\overline{BE}, \tag{27}$$

$$X_H = \overline{BC}\cos\beta = \pm\frac{R_j}{R_j + R_S}\overline{CE}, \tag{28}$$

$$Y_H = \overline{BC}\sin\beta = \frac{R_j}{R_j + R_S}\overline{BE}, \tag{29}$$

where the signs of X_G, X_H, respectively, are

$$\begin{cases} + & \text{if} \quad \alpha > \dfrac{\pi}{2}, \\ - & \text{if} \quad \alpha < \dfrac{\pi}{2}, \end{cases} \tag{30}$$

$$\begin{cases} + & \text{if} \quad \beta < \dfrac{\pi}{2}, \\ - & \text{if} \quad \beta > \dfrac{\pi}{2}. \end{cases} \tag{31}$$

Now we have to find a criterion to establish if a point is lying inside the thorus. Considering Figure 3 and applying the cosines' law to the triangle

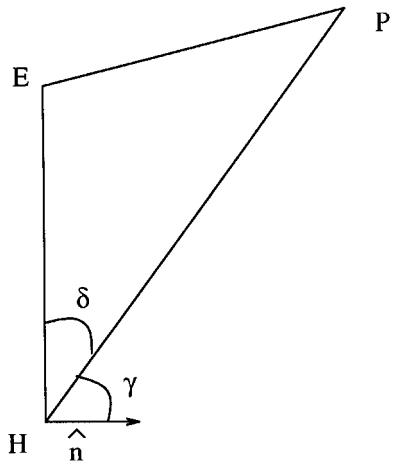

FIGURE 3. Geometrical construction used to determine if the point P lies inside or outside a thorus.

EPH we have

$$|\vec{p}_E - \vec{p}_P| = \sqrt{a_{\tau_j}^2 + |\vec{p}_P - \vec{p}_{\tau_{ij}}|^2 - 2a_{\tau_{ij}}|\vec{p}_P - \vec{p}_{\tau_{ij}}|\cos\delta}, \quad (32)$$

$$\cos\delta = \sin\gamma = \sqrt{1 - \cos^2\gamma}, \quad (33)$$

$$\cos\gamma = \hat{n}_{\tau_{ij}} \cdot \frac{\vec{p}_P - \vec{p}_{\tau_{ij}}}{|\vec{p}_P - \vec{p}_{\tau_{ij}}|}. \quad (34)$$

The point P lies inside the thorus τ_{ij} if

$$|\vec{p}_E - \vec{p}_P| < R_S. \quad (35)$$

We can now define the term $\mathscr{M}_{\mathscr{S}_e}^{ij}(\vec{p})$:

$$\mathscr{M}_{\mathscr{S}_e}^{ij}(\vec{p}) = \begin{cases} 1 & \text{if } \vec{p} \text{ lies inside } \omega_{ij} \text{ but does} \\ & \quad \text{not lie inside } \tau_{ij}, \\ 0 & \text{otherwise.} \end{cases} \quad (36)$$

$\mathscr{M}_{\mathscr{S}_e}^{ij}(\vec{p})$ is also not equal to zero in some regions of space not included in the formal definition but in which $\mathscr{M}_{\mathscr{S}_e}^{i}(\vec{p})$ and/or $\mathscr{M}_{\mathscr{S}_e}^{j}(\vec{p})$ are not equal to zero. This difference in the definition does not affect the value of the molecular shape function $\mathscr{M}_{\mathscr{S}_e}(\vec{p})$ given in terms of OR logical operators.

THREE ATOM TERMS

To define the term $\mathscr{M}_{\mathscr{S}_e}^{ijk}(\vec{p})$ we have to know the positions of the two spheres σ_{ijk}^1 and σ_{ijk}^2 and the radius and the position of the sphere circumscribed to the region in which $\mathscr{M}_{\mathscr{S}_e}^{ijk}(\vec{p}) \neq 0$ but $\mathscr{M}_{\mathscr{S}_e}^{i}(\vec{p}) = \mathscr{M}_{\mathscr{S}_e}^{j}(\vec{p}) = \mathscr{M}_{\mathscr{S}_e}^{k}(\vec{p}) = 0$. We indicate this sphere with θ_{ijk}. Consider now Figure 4: we indicate the positions of the three atoms i, j, k with the letters A, B, C and the positions of the spheres $\sigma_{ijk}^{1,2}$ with D_1 and D_2. The letters $E_{1,2}$, $F_{1,2}$, and $G_{1,2}$ indicate the points of tangency between the atomic spheres and $\sigma_{ijk}^{1,2}$. A simple geometrical consideration leads to the following identities:

$$\begin{aligned} \overline{AD_1} &= \overline{AD_2} = R_i + R_S, \\ \overline{BD_1} &= \overline{BD_2} = R_j + R_S, \\ \overline{CD_1} &= \overline{CD_2} = R_k + R_S; \end{aligned} \quad (37)$$

thus, the problem of finding $\vec{p}_{\sigma_{ijk}^{1,2}}$ reduces to that of finding the intersection of three spheres centered in A, B, C with radii $R_i + R_S$, $R_j + R_S$, $R_k + R_S$, respectively:

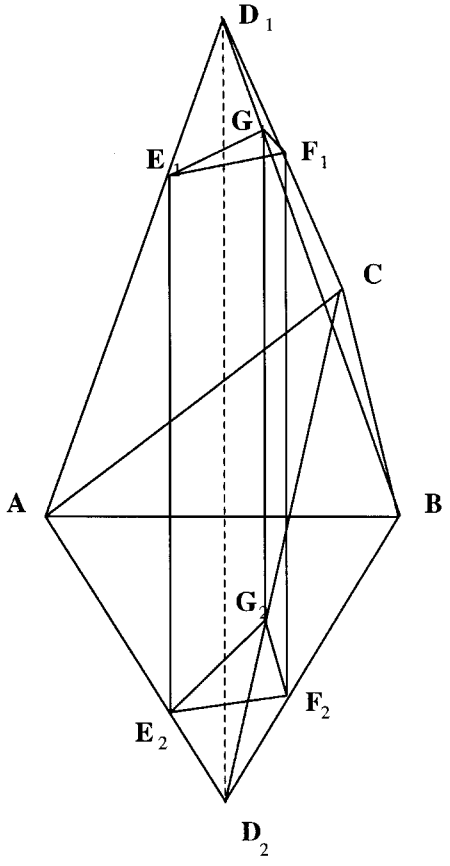

FIGURE 4. Geometrical construction to determine the parameters of the three atom terms.

$$\begin{cases} |\vec{p}_{\sigma_{ijk}^{1,2}} - \vec{p}_i| = R_i + R_S, \\ |\vec{p}_{\sigma_{ijk}^{1,2}} - \vec{p}_j| = R_j + R_S, \\ |\vec{p}_{\sigma_{ijk}^{1,2}} - \vec{p}_k| = R_k + R_S. \end{cases} \quad (38)$$

The solution of this system was given by Dodd and Theodurou.[21]

To find $\vec{p}_{\theta_{ijk}}$ and $R_{\theta_{ijk}}$ observe that the points $E_{1,2}$, $F_{1,2}$, $G_{1,2}$ lie on the surface of θ_{ijk}. This leads to the system

$$\begin{cases} |\vec{p}_{\theta_{ijk}} - \vec{p}_{E_{1,2}}| = R_{\theta_{ijk}}, \\ |\vec{p}_{\theta_{ijk}} - \vec{p}_{F_{1,2}}| = R_{\theta_{ijk}}, \\ |\vec{p}_{\theta_{ijk}} - \vec{p}_{G_{1,2}}| = R_{\theta_{ijk}}. \end{cases} \quad (39)$$

Given the symmetry of systems (38) and (39) with respect to the plane that contains the three atoms i, j, k, only four of these six equations are

linearly independent; thus, a solution always exists. Now we only need the coordinates of the six points of tangency. We remark that

$$\begin{aligned}\overline{AE_1} &= R_i,\\ \overline{D_1E_1} &= R_S;\end{aligned} \tag{40}$$

thus,

$$\vec{p}_{E_1} = \vec{p}_i + \frac{R_i}{R_i + R_S}\left(\vec{p}_{\theta_{ijk}} - \vec{p}_i\right). \tag{41}$$

We can write similar equations for the remaining five points.

The three atom term $\mathscr{M}_{\mathscr{S}_e}^{ijk}(\vec{p})$ is defined as

$$\mathscr{M}_{\mathscr{S}_e}^{ijk}(\vec{p}) = \begin{cases} 1 & \text{if } \vec{p} \text{ belongs to } \theta_{ijk} \text{ but does} \\ & \quad \text{not belong to } \sigma_{i_j^{1,2}}, \\ 0 & \text{otherwise.} \end{cases} \tag{42}$$

The discrepancy between the formal and explicit definition of $\mathscr{M}_{\mathscr{S}_e}^{ij}(\vec{p})$ terms exist also for the $\mathscr{M}_{\mathscr{S}_e}^{ijk}(\vec{p})$ terms, and it is always not significant for the same reason expressed earlier.

DefPol Procedure

The DefPol procedure builds a molecular surface by applying successive deformations to a polyhedron inscribed in a sphere. The steps of the procedure are

1. construction of the polyhedron,
2. projection onto the ellipsoid,
3. displacement of vertices,
4. displacement of centers,
5. recursive generation, and
6. tessellation.

For $\mathscr{S}_e$ we perform the first three steps using the preliminary shape function, repeat the third step starting from preliminary vertices, and then perform the last three steps.

CONSTRUCTION OF POLYHEDRON

The DefPol procedure only uses polyhedra with triangular faces inscribed into a sphere. A polyhedron is characterized by:

- its number of vertices N_v, edges N_e, and faces N_t: these quantities are related by the Euler–Poincaré relation

$$N_t + N_v - N_e = 2. \tag{43}$$

- the coordinates of the vertices ($\vec{v}_i$ with $i = 1, N_v$): we associate an index to any vertex.
- the vertices that belong to any triangle: we store this information as a triad of vertices indexes ($\mathscr{T}_j^1, \mathscr{T}_j^2, \mathscr{T}_j^3$ with $j = 1, N_t$). Also, the triangles are associated with an index.
- the first neighbors of every triangle: we store these data as a triad of triangle indexes ($\mathscr{N}_j^1, \mathscr{N}_j^2, \mathscr{N}_j^3$ with $j = 1, N_t$). The triangle j shares with the triangle $\mathscr{N}_j^1$ the vertices $\mathscr{T}_j^1$ and $\mathscr{T}_j^2$, with the triangle $\mathscr{N}_j^2$ the vertices $\mathscr{T}_j^2$ and $\mathscr{T}_j^3$, and with the triangle $\mathscr{N}_j^3$ the vertices $\mathscr{T}_j^1$ and $\mathscr{T}_j^3$.

In this work we use polyhedra derived by equilateral division from a tetrahedron, icosahedron, and pentakisdodecahedron. The equilateral division is a procedure that replaces a triangle with N_d^2 triangles; therefore, we use polyhedra with $4N_d^2$, $20N_d^2$, and $60N_d^2$ faces.[12,22]

We build the polyhedron inscribed in a unitary sphere centered on the center of mass of the molecule. We also translate the molecule to superimpose the center of mass on the origin of the coordinate system.

PROJECTION ON ELLIPSOID

The first deformation of the polyhedron is its projection on the inertia ellipsoid of the molecule magnified by a factor μ. We have to calculate the inertia tensor $\mathbf{I}$ of the molecule and diagonalize it. The eigenvectors $\vec{x}_1$, $\vec{x}_2$, $\vec{x}_3$ of $\mathbf{I}$ are the principal axes of rotation of the molecule: they are collinear with the ellipsoid axes. The eigenvalues I_1, I_2, I_3 of I are related to the half-axes b_1, b_2, b_3 of the ellipsoid by the relation

$$b_i = \sqrt{\frac{5(\mathrm{tr}(\mathbf{I}) - I_i)}{2M}}, \tag{44}$$

where M is the mass of the molecule. The just introduced factor μ must be large enough to have all the shape function inside the ellipsoid: a value between 2 and 3 is generally sufficient.

To project the polyhedron on the ellipsoid we apply the transformation

$$\vec{v}_i \to \mu\Big[b_1\big(\vec{v}_1 \cdot \vec{x}_1\big)\vec{x}_1 + b_2\big(\vec{v}_2 \cdot \vec{x}_2\big)\vec{x}_2 + b_3\big(\vec{v}_3 \cdot \vec{x}_3\big)\vec{x}_3\Big] \quad (45)$$

to any vertex.

DISPLACEMENT OF VERTICES

We displace any vertex along the line connecting its initial position with the origin of the coordinates until it is lying on the surface of the molecule.

We use a numerical procedure to displace the vertices. First we apply the transformation

$$\vec{v}_i \to \vec{v}_i - \vec{p}_c, \quad (46)$$

where

$$\vec{p}_c = -\frac{1}{N_{\text{div}}^{\text{c}}}\vec{v}_i, \quad (47)$$

where $N_{\text{div}}^{\text{c}}$ is the *number of coarse steps* until

$$\mathscr{M}_{\mathscr{S}_x}\big(\vec{v}_i\big) = 1. \quad (48)$$

Then we invert the last transformation

$$\vec{v}_i \to \vec{v}_i + \vec{p}_c. \quad (49)$$

Now $\vec{v}_i$ is lying outside the molecule.

The second step consists of applying a transformation with a small displacement

$$\vec{v}_i \to \vec{v}_i - \vec{p}_f \quad (50)$$

with

$$\vec{p}_f = \frac{1}{N_{\text{div}}^{\text{f}}}\vec{p}_c, \quad (51)$$

where $N_{\text{div}}^{\text{f}}$ is the *number of fine steps* until

$$\mathscr{M}_{\mathscr{S}_x}\big(\vec{v}_i\big) = 1. \quad (52)$$

We set

$$\vec{v}_i^{+} = \vec{v}_i \quad (53)$$

and

$$\vec{v}_i^{-} = \vec{v}_i + \vec{p}_c \quad (54)$$

and we repeat the bisection procedure

$$\text{if} \quad \mathscr{M}_{\mathscr{S}_x}\left(\frac{\vec{v}_i^{-} + \vec{v}_i^{+}}{2}\right) = \begin{cases} 0 & \vec{v}_i^{-} = \dfrac{\vec{v}_i^{-} + \vec{v}_i^{+}}{2}, \\ 1 & \vec{v}_i^{+} = \dfrac{\vec{v}_i^{-} + \vec{v}_i^{+}}{2}, \end{cases} \quad (55)$$

until

$$|\vec{v}_i^{-} - \vec{v}_i^{+}| < \delta, \quad (56)$$

where δ is a given threshold of convergence.

We finally set

$$\vec{v}_i = \frac{\vec{v}_i^{-} - \vec{v}_i^{+}}{2}. \quad (57)$$

When we have performed this procedure on all the vertices we have a corrugated polyhedron that is topologically equivalent to the original one.

DISPLACEMENT OF CENTERS

We displace any center of the triangular faces along the vector orthogonal to the triangle.

The center of triangle j is

$$\vec{c}_j = \frac{1}{3}\sum_{l=1}^{3}\vec{v}_{\tau_j^l}. \quad (58)$$

If the value of the shape function at $\vec{c}_j$ is 1, then we displace the center along the outward orthogonal vector; otherwise we displace it along the inward orthogonal vector. We now define a triangle index that will be useful later:

$$\iota_j = \begin{cases} 1 & \text{if we move outward,} \\ -1 & \text{if we move inward.} \end{cases} \quad (59)$$

The displacement procedure is similar to that used for the vertices. The unit orthogonal vector is

$$\hat{o} = \pm\frac{\left(\vec{v}_{\mathscr{T}_j^2} - \vec{v}_{\mathscr{T}_j^1}\right) \times \left(\vec{v}_{\mathscr{T}_j^3} - \vec{v}_{\mathscr{T}_j^1}\right)}{\left|\left(\vec{v}_{\mathscr{T}_j^2} - \vec{v}_{\mathscr{T}_j^1}\right) \times \left(\vec{v}_{\mathscr{T}_j^3} - \vec{v}_{\mathscr{T}_j^1}\right)\right|}, \quad (60)$$

where the sign is determined by the displacement direction. The coarse step is defined as

$$\vec{p}_g = \frac{\sqrt{A_i}}{N_{\text{div}}^{\text{c}}}\hat{o}, \quad (61)$$

where A_j is the triangle area. (The step length thus takes account of the triangle size.)

The fine step and the bisection procedure are the same as for the vertices.

RECURSIVE GENERATION

We now consider for every triangle j the tetrahedron with vertices $\vec{c}_j$, $\vec{v}_{\mathcal{T}_j^1}$, $\vec{v}_{\mathcal{T}_j^2}$, and $\vec{v}_{\mathcal{T}_j^3}$. The volume of this solid is

$$V_j = \frac{1}{3} A_j |\vec{c}_j - \vec{c}_{0j}|, \tag{62}$$

where $\vec{c}_{0j}$ is the undisplaced center.

If V_j is larger than a given threshold V_{lim} we proceed to the triangle division. In fact some regions of the molecule could require a more detailed description of the surface.

The procedure involves these transformations on the polyhedron indices:

$$N_\text{v} \to N_\text{v} + 1, \tag{63}$$

$$N_t \to N_t + 2, \tag{64}$$

$$N_\text{e} \to N_\text{e} + 3. \tag{65}$$

We define the new vertex

$$\vec{v}_{N_\text{v}} = c_j, \tag{66}$$

the new triangles

$$\begin{aligned} \mathcal{T}^1_{N_t-1} &= \mathcal{T}^1_j \quad \mathcal{T}^2_{N_t-1} = \mathcal{T}^2_j \quad \mathcal{T}^3_{N_t-1} = N_\text{v}, \\ \mathcal{T}^1_{N_t} &= \mathcal{T}^1_j \quad \mathcal{T}^2_{N_t} = \mathcal{T}^3_j \quad \mathcal{T}^3_{N_t} = N_\text{v}, \\ \mathcal{T}^1_{N_t+1} &= \mathcal{T}^2_j \quad \mathcal{T}^2_{N_t+1} = \mathcal{T}^3_j \quad \mathcal{T}^3_{N_t+1} = N_\text{v}, \end{aligned} \tag{67}$$

and the new neighbors

$$\begin{aligned} \mathcal{N}^1_{N_t-1} &= \mathcal{N}^1_j \quad \mathcal{N}^2_{N_t-1} = N_t \quad \mathcal{N}^3_{N_t-1} = N_{t+1}, \\ \mathcal{N}^1_{N_t} &= \mathcal{N}^2_j \quad \mathcal{N}^2_{N_t} = N_{t-1} \quad \mathcal{N}^3_{N_t} = N_{t+1}, \\ \mathcal{N}^1_{N_t+1} &= \mathcal{N}^3_j \quad \mathcal{N}^2_{N_t+1} = N_{t-1} \quad \mathcal{N}^3_{N_t+1} = N_t. \end{aligned} \tag{68}$$

The triangle j is no further referenced. We recursively apply this procedure to the new triangles until there are no more triangles to divide. A control of the maximum number of recursive steps has been introduced to avoid an excessive proliferation of triangles.

DEFINITION OF TESSELLATION

The recursive generation gives us a deformed polyhedron with metrics and topology different from the original one. As a final step we transform this polyhedron composed of flat triangles into a polyhedron composed of spherical triangles. For any tesserae we have three vertices and a displaced center: there is only one sphere on which all these four points are lying; thus, to any triangle we associate a sphere. To find the center position $\vec{r}_k$ and the radii R_k of the sphere associated with the triangle k we have to solve the system

$$\begin{cases} |\vec{c}_k - \vec{r}_k| = R_k, \\ |\vec{v}_{\mathcal{T}_k^1} - \vec{r}_k| = R_k, \\ |\vec{v}_{\mathcal{T}_k^2} - \vec{r}_k| = R_k, \\ |\vec{v}_{\mathcal{T}_k^3} - \vec{r}_k| = R_k. \end{cases} \tag{69}$$

From the square of the first equation of the previous system,

$$|\vec{c}_k - \vec{r}_k|^2 = R_k^2, \tag{70}$$

$$|\vec{c}_k|^2 + |\vec{r}_k|^2 - 2\vec{r}_k \cdot \vec{c}_k = R_k^2; \tag{71}$$

taking

$$k = R_k^2 - |\vec{r}_k|^2 \tag{72}$$

we have

$$2\vec{r}_k \cdot \vec{c}_k + k = |\vec{c}_k|^2; \tag{73}$$

thus, the quadratic system of eq. (69) can be reduced to a linear system.

This transformation does not affect the connectivity of the triangles: if two flat triangles share an edge the corresponding spherical triangles share an arc. Spherical triangles and flat triangles share the same vertices.

Given the center and the radius of all of the associated spheres, we can calculate the equations of the arcs shared between the neighbor triangles and the area of each spherical triangle. We use the procedure reported by Dodd and Theodorou.[21]

We can define an outward unit vector

$$\hat{n}_k = \frac{\vec{v}_{\text{cal}\,T_k^1} + \vec{v}_{\text{cal}\,T_k^2} + \vec{v}_{\text{cal}\,T_k^2} - 3\vec{r}_k}{\left|\vec{v}_{\text{cal}\,T_k^1} + \vec{v}_{\text{cal}\,T_k^2} + \vec{v}_{\text{cal}\,T_k^2} - 3\vec{r}_k\right|} \iota_k \tag{74}$$

and an average point

$$\vec{s}_k = \vec{r}_k + R_k \hat{n}_k. \tag{75}$$

We have now for any tessera an analytical area, an outward unit vector, and an average point: this mathematical structure is a surface tessellation. A tessellation allows us to perform numerical integrations and to solve BEM problems on the surface. We can also easily compute the molecular volume by applying the Gauss theorem.

IMPLEMENTATION

In this section we give some implementation details regarding an efficient evaluation of the shape function.

We can safely neglect some small terms of the shape function. A large number of the two and three atom terms are not equal to zero in a very small portion of space. We neglect the two atom term $\mathcal{M}^{ij}_{\mathcal{S}_e}(\vec{p})$ if

$$|\vec{r}_j - \vec{r}_j| < R_> + R_< (1 - 2\varrho), \tag{76}$$

where

$$R_> = \mathrm{Max}(R_i, R_j), \tag{77}$$

$$R_< = \mathrm{Min}(R_i, R_j), \tag{78}$$

and ϱ is a parameter given as input. This test controls the overlap of the two atomic spheres so that it is not too large: in this case the two atom term is not significant. The value of ϱ lies in the range 0–1: when we increase ϱ we increase the number of two atom terms.

This test is also applied to the circumscribed speres associated with two and three atom terms. If one of these spheres exhibits a too large overlap with an atomic sphere or another circumscribed sphere, the corresponding term is neglected. This test eliminates redundant or nonsignificant terms in some cases of local symmetry (e.g., three or more aligned atoms, six-member rings, etc.).

Furthermore, a term $\mathcal{M}^{ijk}_{\mathcal{S}_e}(\vec{p})$ is neglected if one or more of the terms $\mathcal{M}^{ij}_{\mathcal{S}_e}(\vec{p})$, $\mathcal{M}^{ik}_{\mathcal{S}_e}(\vec{p})$, $\mathcal{M}^{jk}_{\mathcal{S}_e}(\vec{p})$ is neglected.

In the DefPol definition of $\mathcal{S}_e$ a two or three atom term of the shape function is significant only if it is lying on the periphery of the portion of space where the shape function is not equal to zero. Thus, we have several terms that are embedded in the bulk of the molecule and are not significant. We found it convenient to perform a preliminary calculation using only monoatomic terms with the set of radii used for $\mathcal{S}_e$ increased by R_s. This preliminary cavity includes the $\mathcal{S}_e$ cavity. Looking at the definition of the two and three atoms terms, it is evident that only the atoms lying on this surface are significant for the construction of $\mathcal{S}_e$, so the $\mathcal{S}_e$ shape function is built by only using these atoms.

Numerical Results

We present here a comparison between DefPol and GEPOL results. Both sets of results were obtained by inserting the cavity subroutines in the same general program for solvation effects, namely, PCM.

There are now numerous versions of PCM[17]: we use the standard BEM-PCM version[23] with a further implementation that also allows nonquantum calculations of the electrostatic contribution to the solvation free energy ΔG_{el}. We recall that in GEPOL we have only convex tesserae but in DefPol they may also be concave. This fact affects the sign of the curvature correction factor in the diagonal term of the PCM-BEM matrix.[24] The ι_i factor defined in eq. (59), which is 1 if the tesserae is convex and -1 if it is concave, corresponds to this sign.

Table I presents the values of the area of the three kinds of surfaces, $\mathcal{S}_w$, $\mathcal{S}_e$, and $\mathcal{S}_a$. We recall that we used the original van der Waals radii (Bondi compilation[25]) to define the shape function $\mathcal{M}_{\mathcal{S}_w}(\vec{p})$, $\mathcal{M}_{\mathcal{S}_a}(\vec{p})$, while we use the same set of radii multiplied by a factor 1.2 to define $\mathcal{M}_{\mathcal{S}_e}(\vec{p})$. The radius of the solvent R_s is set equal to 1.385 Å. Solute geometries were optimized *in vacuo* using the Discover 2.0 program[26] with the native force field. GEPOL gives exact values for $\mathcal{S}_w$ and $\mathcal{S}_a$ (within the choice of cavity radii) because we use the GEPOL-GB version[18] in which there are no truncations for the irregularly shaped tesserae; this procedure also gives results of good quality for $\mathcal{S}_e$ because we used the values of GEPOL thresholds giving the maximum accuracy. DefPol values are all approximate by definition: moreover, we used a large mesh partition of the original sphere as the value that will probably be used in routine calculations on large molecules: by using a finer partition the difference between DefPol and GEPOL surfaces is halved. A graphical comparison between the two procedures is given in Figure 5. The results of a least square regression analysis using the function

$$S_{x,\mathrm{GEPOL}} = aS_{x,\mathrm{DefPol}} \tag{79}$$

is reported in Table II. For brevity we do not report comparisons of the volumes of the cavities: the agreement between GEPOL and DefPol volumes is similar to that shown for the surfaces.

Table III presents the components of the solvation free energy computed at fixed solute geometry for the same set of 80 molecules (see, e.g., Tomasi and Persico[1] for the definition of G_{sol}):

$$G_{sol} = G_{dis} + G_{rep} + G_{cav} + \Delta G_{el}. \tag{80}$$

The ΔG_{el} was computed by using the BEM-PCM procedure at the classical level (i.e., using the

TABLE I.
DefPol-GEPOL Comparison on 80 Neutral Small Molecules at MM level.

	S_a (Å)^2		S_w (Å)^2		S_e (Å)^2	
Molecule	DefPol	GEPOL	DefPol	GEPOL	DefPol	GEPOL
α-*d*-Glucose	341.228	338.796	195.639	200.047	210.556	208.009
β-*d*-Glucose	341.749	340.243	194.126	200.116	209.976	208.834
1,1,1-Trichloroethane	247.941	246.979	118.479	120.262	139.016	139.934
1,1-Dichloroethylene	211.214	210.911	90.624	90.628	113.711	114.266
1,1-Difluoroethane	199.352	199.044	82.764	83.150	100.276	101.302
1,3-Butadiene	221.446	220.499	96.282	97.551	115.368	115.806
1,4-Dibromobenzene	307.799	305.407	153.140	159.206	182.954	181.760
1-Butyne	233.804	229.834	99.373	106.757	125.352	120.324
1-Iodobutane	277.280	275.928	136.793	140.930	159.875	158.652
2-Iodopropane	246.908	246.172	117.810	119.358	137.912	137.413
2-Methylpropylene	232.979	232.492	106.077	105.969	122.623	123.378
3-Hydroxybenzaldehyde	291.670	290.212	144.616	145.023	167.251	167.364
E-dichloroethylene	214.051	212.815	90.981	91.330	114.811	115.264
Z-dichloroethylene	210.889	210.502	89.978	90.355	112.348	112.587
Acetaldehyde	186.375	186.042	73.560	73.906	91.172	91.705
Acetate anion	195.253	194.874	79.685	79.768	98.509	99.118
Acetic acid	218.857	218.636	95.668	96.081	118.249	118.504
Alanine	249.943	248.928	122.362	121.303	142.706	138.665
Aniline	261.638	260.974	123.895	124.374	144.279	144.535
Arginine	378.565	373.179	201.113	204.814	228.664	223.287
Asparagine	289.411	285.853	147.500	152.233	171.170	164.376
Aspartic acid	275.046	273.464	136.226	138.326	156.339	156.458
Benzene	244.103	242.269	109.364	111.567	131.453	130.880
Benzoic acid	294.993	293.115	145.284	147.914	169.274	168.010
Bromoethane	213.106	212.642	91.713	92.078	112.965	111.982
Bromoform	243.237	243.011	112.842	113.123	138.979	138.980
Butane	241.617	240.595	111.942	114.220	129.228	128.007
Butylamine	266.698	263.521	130.276	139.488	147.257	145.272
Butylnitril	281.555	278.521	136.730	141.900	156.587	155.442
Chlorobenzene	267.352	266.637	126.379	127.053	150.520	150.266
Chloroform	222.678	222.190	98.493	98.790	122.267	122.539
Cyanidric acid	141.869	141.575	47.987	48.301	64.956	64.619
Cyclopentene	239.249	238.976	108.291	109.031	127.437	127.782
Cysteine	264.535	263.201	128.570	130.374	148.746	148.641
Cytosine	268.271	267.221	127.075	129.928	150.543	148.979
Diethylsulfur	281.888	280.038	135.308	137.941	160.057	155.412
Dimethylether	198.750	198.425	82.443	82.912	98.852	99.706
Dimethylamine	206.614	206.328	87.766	88.192	103.909	104.757
Ethane	182.165	181.864	70.653	70.875	86.811	86.932
Ethanol	197.819	195.796	80.977	82.658	98.480	97.825
Ethylamine	203.357	202.999	86.755	87.175	105.431	102.546
Ethylene	166.316	166.144	61.621	561.835	78.459	78.513
Fluoroform	175.343	175.113	67.454	67.679	85.777	85.613
Glutamic acid	311.494	307.639	159.463	161.912	181.813	179.130
Glutamine	312.744	309.583	160.347	163.512	183.610	181.522
Glycine	206.839	206.651	87.512	88.287	106.174	106.001
Histidine	336.004	329.022	169.075	178.564	196.116	192.368
Indole	291.041	289.571	141.382	143.118	165.439	165.258
Isoleucine	308.729	306.422	164.598	167.156	180.798	180.096
Leucine	317.345	311.026	167.591	173.652	185.084	182.237
Lysine	346.463	343.037	184.392	190.175	204.796	202.499

(Continued)

TABLE I.
(Continued)

Molecule	S_a (Å)^2		S_w (Å)^2		S_e (Å)^2	
	DefPol	GEPOL	DefPol	GEPOL	DefPol	GEPOL
Methane	141.310	141.254	47.405	47.472	61.379	61.333
Methanol	163.868	163.018	59.736	60.334	76.179	75.441
Methanthiol	181.225	180.993	70.597	70.993	89.180	89.447
Methionine	327.350	325.373	172.553	176.093	192.484	191.408
Methylacetate	238.893	237.523	106.825	107.786	126.357	126.684
Methylamine	171.988	171.745	65.050	65.309	80.911	80.824
Naphthalene	309.377	307.685	155.546	158.421	179.210	179.163
Nitrobenzene	285.681	282.959	137.396	140.789	161.837	161.186
Phenanthrene	369.295	366.756	197.581	201.630	224.210	223.660
Phenol	256.832	255.275	119.590	120.889	141.028	141.423
Phenylalanine	348.490	346.583	187.960	191.132	215.414	207.603
Phosphine	152.111	152.039	53.509	54.661	69.751	69.669
Proline	270.439	269.525	134.719	136.950	153.159	151.676
Propane	215.941	213.675	92.549	97.157	113.525	108.798
Propylammine	236.936	234.949	108.850	109.629	128.017	124.723
Pyrene	386.682	383.661	213.018	216.665	237.897	238.282
Pyridine	234.544	233.767	104.216	104.780	125.460	125.565
Pyrrole	221.451	220.182	94.473	94.846	114.751	114.942
Serine	246.292	245.603	118.163	119.128	135.414	135.041
Tetrachloroethylene	249.139	248.773	116.465	116.653	144.391	143.507
Thiophenol	277.055	275.755	132.112	133.905	156.267	155.842
Threonine	273.467	272.246	137.911	138.358	154.545	154.867
Thymine	286.325	284.751	141.890	143.648	164.804	162.932
Toluene	272.766	271.890	130.691	131.990	151.616	151.736
Trimethylamine	235.411	234.602	109.719	110.296	124.378	126.431
Trypticene	485.488	480.905	276.917	286.222	307.276	305.530
Tyroxine	371.270	363.933	202.096	211.776	231.680	220.479
Valine	281.770	280.515	145.677	147.938	160.604	161.018
Water	120.091	120.105	36.104	36.100	48.626	48.593

atomic charges given by the Discover 2.0 program[26]). This calculation is based on the tessellated $\mathscr{S}_e$ surface.

The G_{dis} and G_{rep} were computed using the Floris–Tomasi[27] procedure. The atom–atom parameters were taken from Huron and Claverie.[28] Both terms were computed using the tessellated $\mathscr{S}_a$ surface.

The G_{cav} was computed using Pierotti formula[29] using the $\mathscr{S}_w$ surface without tessellation. In the past[1] we often used the modification of the Pierotti formula suggested by Lainglet et al.[30] Recent simulation results[31] suggest that the original Pierotti formula is probably better for cavities of irregular shape. Figure 6 is a graphical comparison of the G_{sol} values obtained with DefPol and GEPOL.

The results of a least square regression analysis based on functions similar to those given in eq. (79) are reported in Table IV.

The quantity of the most direct practical interest is the total value of the solvation energy. In Table III we also report the difference δ between the values obtained with the two methods: the largest difference (1 kcal/mol) was found for arginine; the unsigned mean error was remarkably low.

Figures 7 and 8 are pictorial representations of the DefPol surfaces.

We also performed some Hartree–Fock (HF) *ab initio* calculations. For 20 molecules of the first series we also performed the calculation of the electrostatic component of the free energy of solvation at the HF level using the STO-3G and 6.31G** basis sets. These results are reported in Table V. The concordance between GEPOL and DefPol results is similar to that obtained in the classical scheme. According to the variational principle a larger basis set leads to a larger ΔG_{el}; thus, the 6.31** values are larger (negative) than the STO-3G

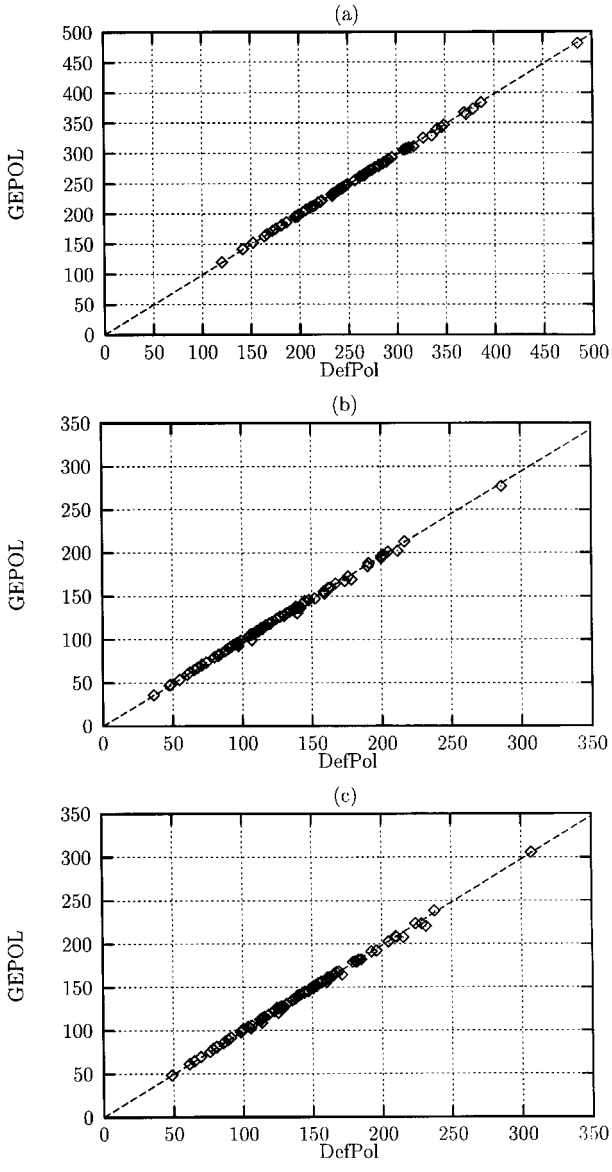

FIGURE 5. Graphical synthesis of Tables I and II. Values for (a) $\mathscr{S}_a$, (b) $\mathscr{S}_w$, and (c) $\mathscr{S}_e$; the points represent the values in Table I, the lines the regression results.

ones. The classic results are closer to the STO-3G results, but there is not a regular trend.

A last series of calculations was carried out for the first eight oligomers of alanine in the α-helix conformation. For the higher terms this corresponds to about two loops of the α helix. The results regarding the calculation of electrostatic [at the molecular mechanics (MM) and HF/STO-3G levels] and nonelectrostatic components of the free energy of solvation, $\mathscr{S}_a$, $\mathscr{S}_w$, $\mathscr{S}_e$, and the number of tesserae involved in $\mathscr{S}_e$ evaluation are reported in Table VI. We note that the GEPOL values are reported only for the first four terms of the series. This is because the GEPOL calculations lead to a number of tesserae higher than the DefPol ones and therefore excessively increase the computational cost (especially memory occupation). Figure 9 reports the number of total and nonempty tesserae generated by GEPOL and the number of tesserae generated by DefPol. It is evident that GEPOL generates a large number of additional spheres that are to a large extent eliminated when GEPOL defines the cavity surface. The same considerations hold for the tesserae of the various spheres. We also note that the fraction of empty tesserae of the total number of tesserae increases as the dimensions of the system increases, especially in systems like an α helix where some nonbonded atoms lie at a distance that allows solvent-obstruction effects. In DefPol we can tune the number of tesserae according to molecular size and avoid the time-wasting generation of nonsignificant tesserae.

TABLE II.
Regression Analysis on Data of Table I.

	r	RMS	*a*
$\mathscr{S}_a$	0.999988	5.81	0.993162
$\mathscr{S}_w$	0.999854	5.87	0.978465
$\mathscr{S}_e$	0.999912	4.33	0.991691

In Table VI we report three different values for ΔG_{el}. The first (MM) is the rigid-charge approximation; the second and third values report two different quantities labeled by QM(x). In the first case (x = U, the label U stands for unpolarized) we simply used the *in vacuo* molecular electrostatic field, disregarding the influence of the reaction field on the molecular wave function. In practice we used the self-consistent field *in vacuo* results to compute ΔG_{el}. In the second case (x = P, the label P stands for polarized) we performed a self-consistent reaction field calculation: the wave function (thus, the molecular electrostatic field) was variationally optimized in the presence of a reaction field. If we compare these quantities with the classical ones we note that

$$|\Delta G_{el}(\mathrm{MM})| < |\Delta G_{el}(\mathrm{QM(U)})| < |\Delta G_{el}(\mathrm{QM(P)})|. \quad (81)$$

The quantum mechanics (unpolarized) [QM(U)] value is closer to MM than QM(P) to QM(U) because in the first two cases the approach is similar:

TABLE III.
DefPol-GEPOL Comparison on 80 Neutral Small Molecules at MM Level for G_{sol} and Its Components (kcal / mol).

	G_{dis}		G_{rep}		G_{cav}		ΔG_{el}		G_{sol}		
Molecule	DefPol	GEPOL	DefPol	GEPOL	DefPol	GEPOL	DefPol	GEPOL	DefPol	GEPOL	$\lvert\delta\rvert$
1,1,1-Trichloroethane	−15.46	−15.56	2.68	2.81	12.25	12.29	−1.72	−1.78	−2.25	−2.24	0.01
1,3-Butadiene	−11.98	−12.02	1.70	1.77	10.99	11.15	−1.24	−1.25	−0.53	−0.35	0.18
1,4-Dibromobenzene	−19.91	−20.17	3.30	3.45	15.62	15.44	−0.75	−0.75	−1.74	−2.03	0.29
1-Butyne	−11.98	−11.94	1.78	1.82	10.01	10.08	−2.77	−2.63	−2.96	−2.67	0.29
1-Iodobutane	−18.12	−18.26	2.74	2.87	14.04	13.96	−0.24	−0.26	−1.58	−1.69	0.11
2-Iodopropane	−15.92	−16.10	2.48	2.61	12.17	12.23	−0.17	−0.18	−1.44	−1.44	0.00
2-Methylpropylene	−11.44	−11.20	1.73	1.76	11.06	10.53	−1.48	−1.22	−0.13	−0.13	0.00
3-Hydroxybenzaldehyde	−14.64	−14.66	2.06	2.11	12.30	12.39	−3.34	−3.37	−3.62	−3.53	0.09
Acetaldehyde	−11.13	−11.19	1.71	1.79	10.12	10.19	−6.50	−6.61	−5.80	−5.82	0.02
Acetate anion	−8.85	−8.82	1.37	1.45	8.14	8.12	−72.40	−72.23	−71.74	−71.48	0.24
Acetic acid	−10.92	−10.79	1.53	1.59	10.21	9.90	−6.73	−6.62	−5.91	−5.92	0.01
Alanine	−14.16	−13.65	2.11	1.99	12.65	12.61	−5.55	−5.55	−4.95	−4.56	0.01
Aniline	−16.90	−16.92	2.37	2.43	14.65	14.73	−3.59	−3.65	−3.47	−3.41	0.06
Arginine	−23.58	−23.40	3.20	3.23	19.53	19.76	−12.47	−11.91	−13.32	−12.32	1.00
Asparagine	−16.51	−16.37	2.31	2.35	15.02	14.93	−11.81	−12.16	−10.99	−11.25	0.26
Aspartic acid	−15.89	−15.89	2.21	2.28	13.82	13.91	−10.62	−10.75	−10.48	−10.45	0.03
Benzene	−15.15	−15.28	2.13	2.21	12.60	12.79	−2.44	−2.44	−2.86	−2.72	0.14
Benzoic acid	−16.36	−16.37	2.18	2.24	14.65	14.77	−6.05	−6.02	−5.58	−5.38	0.20
Bromoethane	−12.03	−12.11	1.98	2.09	9.76	9.82	−1.44	−1.51	−1.73	−1.71	0.02
Bromoform	−16.21	−16.54	3.05	3.24	11.62	11.77	−2.94	−2.85	−4.48	−4.38	0.10
Butane	−11.22	−11.23	1.65	1.72	10.25	10.24	−0.73	−0.75	−0.05	−0.02	0.03
Butylamine	−30.18	−30.26	3.94	3.99	26.42	26.49	−3.81	−3.64	−3.63	−3.42	0.21
Butylnitril	−14.76	−14.52	2.00	2.04	13.92	13.37	−1.78	−1.78	−0.62	−0.89	0.27
1,1-Dichloroethylene	−12.69	−12.83	2.25	2.38	9.64	9.72	−0.19	−0.19	−0.99	−0.92	0.07
1,1-Difluoroethane	−10.08	−10.09	1.52	1.60	8.97	8.98	−3.26	−3.34	−2.85	−2.85	0.00
Chlorobenzene	−15.78	−15.95	2.43	2.54	12.84	13.01	−1.03	−1.02	−1.54	−1.42	0.12
Chloroform	−13.86	−14.01	2.62	2.78	10.36	10.45	−3.49	−3.33	−4.37	−4.11	0.26
Cyanidric acid	−5.61	−5.55	1.08	1.17	5.81	5.67	−1.38	−1.36	−0.10	−0.07	0.03
Cyclopentene	−13.05	−13.02	1.78	1.85	11.72	11.69	−0.24	−0.26	0.21	0.26	0.05
Cysteine	−16.12	−16.17	2.40	2.48	13.13	13.21	−5.90	−5.94	−6.49	−6.42	0.07
Cytosine	−12.43	−12.44	1.73	1.80	11.37	11.48	−3.67	−3.62	−3.00	−2.78	0.22
E-dichloroethylene	−13.02	−13.08	2.36	2.48	9.70	9.75	−0.23	−0.23	−1.19	−1.08	0.11
Z-dichloroethylene	−12.73	−12.92	2.26	2.41	9.61	9.66	−0.40	−0.39	−1.26	−1.24	0.02
Diethylsulfide	−16.42	−16.49	2.28	2.37	13.78	13.83	−0.26	−0.27	−0.62	−0.56	0.06
α-*d*-Glucose	−21.15	−21.06	2.63	2.68	19.12	19.27	−7.49	−7.70	−6.89	−6.81	0.08
β-*d*-Glucose	−21.25	−21.28	2.65	2.72	19.13	19.13	−7.48	−7.54	−6.95	−6.97	0.02
Dimethylether	−9.96	−9.97	1.47	1.54	8.95	8.95	−0.32	−0.33	0.14	0.19	0.05
Dimethylamine	−14.95	−14.81	2.16	2.23	14.13	13.96	−0.71	−0.72	0.63	0.66	0.03
Ethane	−12.89	−12.95	1.75	1.83	11.26	11.35	−1.25	−1.34	−1.13	−1.11	0.02
Ethanol	−10.03	−9.83	1.46	1.50	8.93	8.82	−2.34	−2.30	−1.98	−1.81	0.17
Ethylamine	−10.45	−10.46	1.52	1.59	9.42	9.45	−3.05	−3.06	−2.56	−2.48	0.08
Ethylene	−8.64	−8.60	1.29	1.36	7.87	7.85	−0.23	−0.24	0.29	0.37	0.08
Fluoroform	−8.67	−8.66	1.43	1.51	7.58	7.54	−3.56	−3.57	−3.22	−3.18	0.04
Glutamic acid	−18.00	−17.96	2.46	2.52	15.86	16.02	−9.61	−9.71	−9.29	−9.13	0.16
Glutamine	−18.77	−18.66	2.58	2.62	15.99	16.10	−9.75	−9.94	−9.95	−9.88	0.07
Glycine	−10.68	−10.73	1.61	1.69	9.43	9.43	−7.59	−7.62	−7.23	−7.23	0.00
Histidine	−19.66	−19.41	2.74	2.76	17.29	16.88	−8.37	−8.10	−8.00	−7.87	0.13
Indole	−16.88	−16.92	2.37	2.43	14.28	14.42	−5.30	−5.40	−5.53	−5.47	0.06
Isoleucine	−18.38	−18.32	2.42	2.46	16.31	16.48	−6.22	−6.15	−5.87	−5.53	0.34
Leucine	−18.55	−18.22	2.45	2.45	16.87	16.75	−6.61	−6.57	−5.84	−5.59	0.25
Lysine	−20.63	−20.45	2.73	2.76	18.28	18.26	−7.93	−7.85	−7.55	−7.28	0.27

(Continued)

TABLE III.
(Continued)

Molecule	G_{dis}		G_{rep}		G_{cav}		ΔG_{el}		G_{sol}		$\lvert\delta\rvert$
	DefPol	GEPOL	DefPol	GEPOL	DefPol	GEPOL	DefPol	GEPOL	DefPol	GEPOL	
Methane	−7.49	−7.48	1.19	1.27	7.05	6.99	−0.32	−0.34	0.43	0.44	0.01
Methanol	−7.65	−7.54	1.20	1.26	6.91	6.81	−2.47	−2.47	−2.47	−2.01	0.36
Methanthiol	−10.27	−10.34	1.73	1.83	7.88	7.84	−0.45	−0.45	−1.11	−1.12	0.01
Methionine	−20.08	−20.17	2.77	2.86	17.08	17.20	−6.51	−6.47	−6.74	−6.58	0.16
Methylacetate	−12.50	−12.45	1.82	1.88	11.15	11.22	−2.44	−2.56	−1.97	−1.91	0.06
Methylamine	−10.39	−10.37	1.50	1.58	9.33	9.36	−1.55	−1.67	−1.11	−1.10	0.01
Naphthalene	−16.90	−16.97	2.37	2.43	14.40	14.67	−6.93	−6.94	−7.06	−6.81	0.25
Nitrobenzene	−18.69	−18.78	2.53	2.57	15.56	15.66	−2.05	−1.99	−2.65	−2.54	0.11
Phenanthrene	−13.52	−13.48	1.93	1.97	11.48	11.45	−1.36	−1.31	−1.47	−1.37	0.10
Phenol	−23.75	−23.86	3.12	3.16	19.26	19.44	−2.69	−2.59	−4.06	−3.85	0.21
Phenylalanine	−20.70	−20.81	2.75	2.83	18.36	18.58	−8.17	−7.69	−7.76	−7.09	0.67
Phosphine	−8.22	−8.26	1.56	1.66	6.39	6.21	−0.42	−0.44	−0.69	−0.83	0.14
Proline	−15.25	−15.26	2.02	2.09	13.70	13.77	−8.31	−8.05	−7.84	−7.45	0.39
Propane	−6.13	−6.08	1.00	1.08	5.73	5.62	−0.07	−0.08	0.53	0.54	0.01
Propylammine	−8.19	−8.15	1.27	1.34	7.37	7.31	−1.87	−1.89	−1.42	−1.39	0.03
Pyrene	−16.26	−16.15	2.26	2.31	14.03	14.02	−0.99	−0.97	−0.96	−0.79	0.17
Pyridine	−17.24	−17.33	2.38	2.44	14.23	14.38	−2.97	−2.84	−3.60	−3.35	0.25
Pyrrole	−13.02	−13.08	1.92	1.98	10.89	10.98	−1.77	−1.72	−1.98	−1.84	0.14
Serine	−13.62	−13.67	1.91	1.99	12.15	12.26	−7.49	−7.69	−7.05	−7.11	0.06
Tetrachloroethylene	−17.21	−17.53	3.15	3.32	11.93	12.10	−0.00	−0.00	−2.13	−2.11	0.02
Thiophenol	−25.43	−25.53	3.30	3.32	20.54	20.82	−2.89	−2.77	−4.48	−4.16	0.32
Threonine	−15.58	−15.59	2.11	2.17	13.82	14.06	−8.22	−8.21	−7.87	−7.57	0.30
Thymine	−15.65	−15.73	2.30	2.37	13.09	13.08	−5.01	−5.08	−5.27	−5.36	0.09
Toluene	−16.98	−17.07	2.54	2.61	13.43	13.54	−1.35	−1.32	−2.36	−2.24	0.12
Trimethylamine	−12.61	−12.50	1.76	1.81	11.31	11.40	−1.44	−1.57	−0.98	−0.86	0.12
Trypticene	−15.47	−15.57	2.14	2.20	13.27	13.41	−1.35	−1.31	−1.41	−1.27	0.14
Tyroxine	−22.25	−21.98	2.97	2.99	20.12	19.84	−9.73	−9.75	−8.89	−8.90	0.01
Valine	−24.79	−24.40	3.29	3.30	22.34	21.55	−8.26	−8.58	−7.42	−8.13	0.71
Water	−4.81	−4.72	0.83	0.91	4.66	4.51	−5.17	−5.37	−4.49	−4.67	0.18
Unsigned mean error											0.14

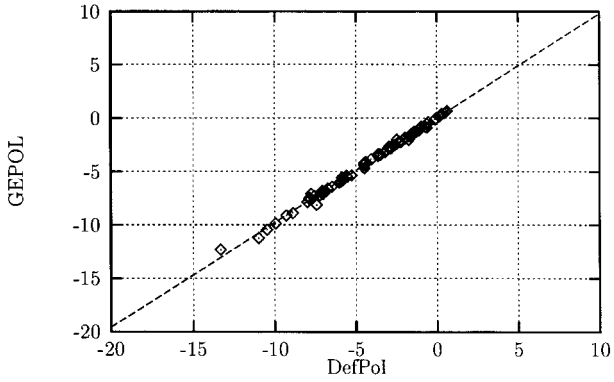

FIGURE 6. Graphical synthesis of Tables III and IV. Only G_{sol} values are represented. The points represent the values in Table III, the lines the regression results.

the ΔG_{el} calculation does not modify the description of the molecule. The residual quantum effect contained in $\Delta G_{el}(MM) - \Delta G_{el}(QM(U))$ is due to the fact that in QM(U) the molecular electrostatics is described by pointlike nuclear charges and by electronic clouds instead of rigid punctual charges alone.

TABLE IV.
Regression Analysis on Data of Table III.

	r	RMS	*a*
G_{dis}	0.99997	0.13	1.000131
G_{rep}	0.99985	0.04	1.031305
G_{cav}	0.99991	0.18	1.000101
ΔG_{el}	0.99962	0.14	0.999995
G_{sol}	0.99990	0.20	0.979053

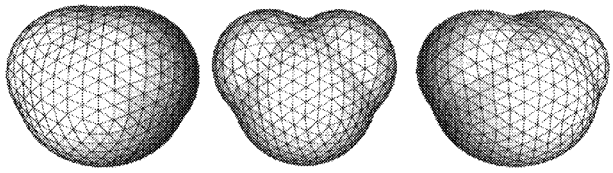

FIGURE 7. From left to right; the $\mathscr{S}_a$, $\mathscr{S}_w$, and $\mathscr{S}_e$ for water.

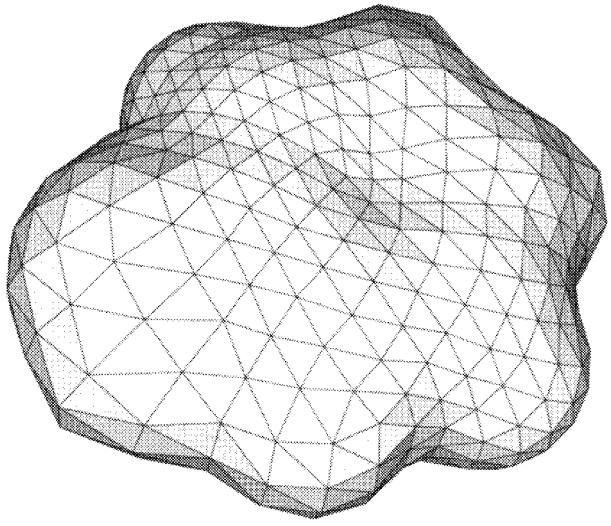

FIGURE 8. $\mathscr{S}_e$ for phenol; the OH group is at the left of the cavity.

Conclusions

The capacity of limiting the number of tesserae to those required for convergence of the calculation of thermodynamical quantities is an important feature. The PCM-BEM procedure requires an amount of memory proportional to N_t^2 and a CPU time proportional to N_t^3. These requirements limit the number of tesserae that we can handle on a computer. The most important limitation at the present state of art of computing is the finite memory size: we can theoretically perform a calculation for an indefinite amount of time on a computer; we are able to stop and restart calculations, but the amount of memory is a physical limit that we can remove only with inefficient memory-disk swapping techniques.

DefPol introduces the possibility of controlling the value of N_t, but it does not influence the

TABLE V.
DefPol-GEPOL Comparison on 20 Small Molecules at HF *Ab Initio* Level for ΔG_{el} (kcal / mol).

	6.31G**			STO-3G		
	DefPol	GEPOL	$\lvert\delta\rvert$	DefPol	GEPOL	$\lvert\delta\rvert$
Acetic acid	−10.24	−10.14	0.10	−5.46	−5.42	0.04
Butylnitrile	−7.03	−6.96	0.07	−4.34	−4.34	0.00
Chloroform	−3.01	−3.00	0.01	−2.75	−2.75	0.00
Dimethylammine	−4.00	−3.95	0.05	−3.02	−2.96	0.06
Diethylsulfur	−3.38	−3.13	0.25	−1.30	−1.18	0.12
Difluoroethane	−4.38	−4.34	0.04	−1.90	−1.88	0.02
Dimethylether	−4.01	−4.00	0.01	−2.37	−2.36	0.01
Cyanidric acid	−7.00	−6.97	0.03	−3.94	−3.92	0.02
Ethane	−0.13	−0.13	0.00	−0.12	−0.12	0.00
Ethanol	−6.63	−6.34	0.29	−4.01	−3.80	0.21
Methane	−0.12	−0.12	0.00	−0.12	−0.12	0.00
Methanol	−6.77	−6.71	0.06	−4.11	−4.06	0.05
Methylacetate	−7.59	−7.65	0.06	−3.62	−3.70	0.08
Methylamine	−5.46	−5.42	0.04	−4.10	−4.08	0.02
p-Dibromobenzene	−3.81	−3.91	0.10	−1.89	−1.93	0.04
Phenol	−7.28	−7.15	0.13	−4.14	−4.03	0.11
Phospine	−3.31	−3.29	0.02	−1.67	−1.64	0.03
Pyridine	−5.31	−5.35	0.04	−3.14	−3.10	0.04
Trichloroethylene	−1.67	−1.65	0.02	−2.05	−2.03	0.02
Water	−9.10	−8.99	0.01	−5.74	−5.68	0.06
	unsigned mean error		0.07	unsigned mean error		0.05

TABLE VI.
DefPol-GEPOL Comparison at Various Levels of Theory on First Eight Terms of Polyalanine Series.

N	Method	ΔG_{el} MM	ΔG_{el} QM(U)	ΔG_{el} QM(P)	G_{cav}	G_{dis}	G_{rep}	$\mathscr{S}_a$	$\mathscr{S}_w$	$\mathscr{S}_w$	N_T
1	DefPol	−5.55	−6.63	−7.30	12.65	−14.16	2.11	249.943	122.362	142.706	540
	GEPOL	−5.55	−6.74	−7.43	12.61	−13.65	1.99	248.928	121.303	138.665	422 (960)
2	DefPol	−8.75	−10.11	−11.39	20.39	−21.69	3.01	361.206	208.090	230.915	720
	GEPOL	−8.56	−9.94	−11.14	20.31	−20.64	2.82	357.261	203.968	220.687	845 (2520)
3	DefPol	−10.22	−12.97	−14.59	28.82	−28.08	3.76	459.922	303.357	319.128	960
	GEPOL	−9.75	−12.77	−14.34	27.48	−26.74	3.54	456.480	288.115	297.194	1683 (7200)
4	DefPol	−15.69	−17.14	−19.95	35.43	−33.82	4.40	535.649	379.365	398.481	1280
	GEPOL	−15.62	−17.49	−20.10	34.79	−32.03	4.14	531.428	371.483	358.656	2028 (8280)
5	DefPol	−17.13	−18.39	−20.77	44.27	−39.49	5.09	609.012	450.681	461.995	1620
	GEPOL										
6	DefPol	−19.78	−21.32	−23.96	54.53	−45.34	5.80	688.571	536.993	528.213	1620
	GEPOL										
7	DefPol	−22.42	−24.31	−27.34	58.21	−50.80	6.37	769.124	608.683	592.708	2000
	GEPOL										
8	DefPol	−24.74	−27.38	−30.78	63.74	−55.63	6.98	855.555	678.273	664.981	2000
	GEPOL										

memory requirements of the PCM-BEM procedure.

Of course, it is possible to introduce in the computational strategy techniques similar to the direct HF methods now in use. They could be combined with other additional techniques for reducing computational times: some hints were offered in a recent review[32]; a procedure addressing the reduction of the N_t^3 scaling time and the N_t^2 scaling memory for the self-equilibration of the ASCs, based on a fast multipole expansion, is now under testing.[33] These modifications of the computational strategy introduce, however, some additional problems in the analytical calculation of the nuclear geometry derivatives. However, we stress that the present formulation can be applied to a remarkably large set of compounds: we performed a G_{sol} calculation for some globular proteins with 3000–4000 atoms at the MM level without difficulty. As an example the HIV protease enzyme was presented in a preliminary collective report on work in progress in our group[34]: with a number of tesserae of the order of 1000–2000 we obtained G_{sol} values that we estimated to have a precision analogous to that of the examples shown here. It is also probably possible to use a similar number of tesserae for larger solutes: in this case it should be necessary to replace the single point BEM procedure used until now in PCM procedures with a more detailed sampling of each tessera. Many point BEM procedures, analytical derivatives, fast multipole expansions, and specialized versions of DefPol for a molecular system with one hole (topologically equivalent to a thorus) or more will be the subject of future reports.

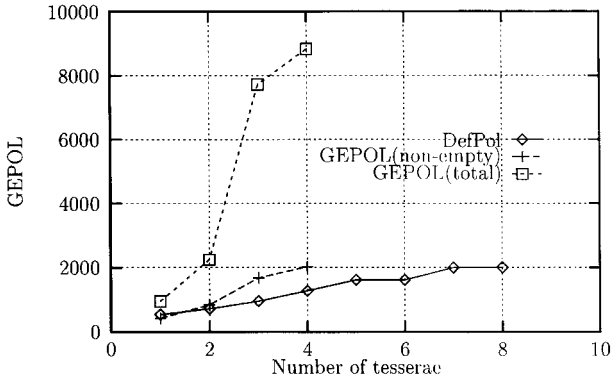

FIGURE 9. The increase of the number of tesserae along the polyalanine series in DefPol and GEPOL.

References

1. J. Tomasi and M. Persico, *Chem. Rev.*, **94**, 2027 (1994).
2. A. Y. Meyer, *Chem. Soc. Rev.*, **446**, 15 (1986).
3. P. G. Mezey, *Rev. Comput. Chem.*, **1**, 265 (1990).
4. See, for example, (a) W. Hasel, T. F. Hendrickson, and W. C. Still, *Tetrahedron Comput. Methodol.* **1**, 103 (1998); (b) G. Perrot, B. Cheng, K. D. Gibson, J. Vila, K. A. Palmer, A.

Nayeem, B. Maigret, and H. A. Scheraga, *J. Comput. Chem.*, **13**, 1 (1992); (c) S. M. Le Grand and J. K. M. Merz, *J. Comput. Chem.*, **14**, 349 (1993); (d) M. Petitjean, *J. Comput. Chem.*, **15**, 507 (1994); (e) D. A. Liotard, G. D. Hawkins, G. C. Lynch, C. J. Cramer, and D. G. Truhlar, *J. Comput. Chem.*, **16**, 422 (1995); (f) A. A. Bliznyuk and J. E. Gready, *J. Comput. Chem.*, **17**, 962 (1996); (g) A. A. Bliznyuk and J. E. Gready, *J. Comput. Chem.*, **17**, 970 (1996); (h) C. E. Kundrot, J. W. Ponder, and F. M. Richards, *J. Comput. Chem.*, **12**, 402 (1991); (i) V. Gogonea and E. Osawa, *J. Mol. Struct.* (*Theochem*), **311**, 305 (1994); (j) V. Gogonea and E. Osawa, *Supramol. Chem.*, **3**, 303 (1994).

5. M. L. Connolly, http://www.netsci.org/Science/Compchem/feature14.html.
6. M. L. Connolly, *J. Am. Chem. Soc.*, **107**, 1118 (1985).
7. M. L. Connolly, *Science*, **221**, 709 (1983).
8. J. L. Pascual Ahuir, E. Silla, J. Tomasi, and R. Bonaccorsi, *J. Comput. Chem.*, **8**, 778 (1987).
9. J. L. Pascual Ahuir, E. Silla, and I. Tuñon, *J. Comput. Chem.*, **15**, 1127 (1994).
10. R. J. Zauhar and H. S. Morgan, *J. Comput. Chem.*, **9**, 171 (1988).
11. R. J. Zauhar and H. S. Morgan, *J. Comput. Chem.*, **11**, 603 (1990).
12. C. S. Pomelli and J. Tomasi, *Theor. Chem. Acc.*, to appear.
13. J. L. Pascual Ahuir and E. Silla, *J. Comput. Chem.*, **11**, 1047 (1990).
14. S. Miertuš, E. Scrocco, and J. Tomasi, *Chem. Phys.*, **55**, 117 (1981).
15. D. Rinaldi, M. F. Ruiz−Lopez, and J. L. Rivail, *J. Chem. Phys.*, **78**, 834 (1983).
16. C. J. Cramer and D. G. Truhlar, *J. Am. Chem. Soc.*, **113**, 9901 (1991).
17. (a) C. Amovilli, V. Barone, R. Cammi, E. Cancès, M. Cossi, B. Mennucci, C. S. Pomelli, and J. Tomasi, *Adv. Quantum Chem.*, to appear; (b) E. Cancès, B. Mennucci and J. Tomasi, *J. Chem. Phys.*, **109**, 260 (1998).
18. M. Cossi, B. Mennucci, and J. Tomasi, *Chem. Phys. Lett.*, **228**, 165 (1994).
19. B. Lee and F. M. Richards, *J. Mol. Biol.*, **55**, 379 (1971).
20. F. M. Richards, *Annu. Rev. Biophys. Bioeng.*, **6**, 151 (1977).
21. L. R. Dodd and D. N. Theodorou, *Mol. Phys.*, **72**, 1313 (1991).
22. M. J. Wenninger, *Spherical Models*, Cambridge University Press, Cambridge, U.K., 1979.
23. R. Cammi and J. Tomasi, *J. Chem. Phys.*, **100**, 7495 (1994).
24. B. Wang and G. P. Ford, *J. Chem. Phys.*, **97**, 4162 (1992).
25. A. Bondi, *J. Phys. Chem.*, **68**, 441 (1964).
26. *Discover 2.5 User Guide*, BIOSYM Technologies, Inc., San Diego, CA, 1989.
27. F. Floris and J. Tomasi, *J. Comput. Chem.*, **5**, 616 (1989).
28. M. J. Huron and P. J. Claverie, *J. Phys. Chem.*, **76**, 2123 (1972).
29. R. A. Pierotti, *J. Phys. Chem.*, **67**, 1840 (1963).
30. J. Lainglet, P. J. Claverie, J. Caillet, and A. Pullman, *J. Phys. Chem.*, **92**, 1617 (1988).
31. F. M. Floris, M. Selmi, A. Tani, and J. Tomasi, *J. Chem. Phys.*, **107**, 6353 (1997).
32. J. Tomasi, B. Mennucci, and R. Cammi, In *Molecular Electrostatic Potentials*: *Concepts and Applications*, J. S. Murray and K. Sen, Eds., Elsevier, New York, 1996, p. 1.
33. C. S. Pomelli, under development.
34. R. Cammi, M. Cossi, B. Mennucci, C. S. Pomelli, and J. Tomasi, *Int. J. Quantum Chem.*, **60**, 1165 (1996).